

Louisiana State University  
**LSU Digital Commons**

---

Faculty Publications

Department of Biological Sciences

---

3-15-2007

## Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree

Jeffery M. Saarela  
*The University of British Columbia*

Hardeep S. Rai  
*The University of British Columbia*

James A. Doyle  
*The University of British Columbia*

Peter K. Endress  
*University of California, Davis*

Sarah Mathews  
*University of Zurich*

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.lsu.edu/biosci\\_pubs](https://digitalcommons.lsu.edu/biosci_pubs)

---

### Recommended Citation

Saarela, J., Rai, H., Doyle, J., Endress, P., Mathews, S., Marchant, A., Briggs, B., & Graham, S. (2007). Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. *Nature*, 446 (7133), 312-315. <https://doi.org/10.1038/nature05612>

This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Digital Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Digital Commons. For more information, please contact [ir@lsu.edu](mailto:ir@lsu.edu).

---

## Authors

Jeffery M. Saarela, Hardeep S. Rai, James A. Doyle, Peter K. Endress, Sarah Mathews, Adam D. Marchant, Barbara G. Briggs, and Sean W. Graham

## LETTERS

# Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree

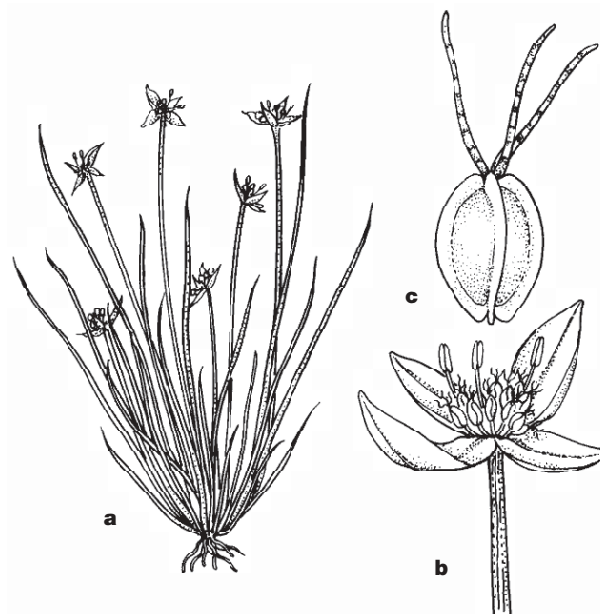
Jeffery M. Saarela<sup>1†</sup>, Hardeep S. Rai<sup>1</sup>, James A. Doyle<sup>2</sup>, Peter K. Endress<sup>3</sup>, Sarah Mathews<sup>4</sup>, Adam D. Marchant<sup>5</sup>, Barbara G. Briggs<sup>5</sup> & Sean W. Graham<sup>1</sup>

Although the relationship of angiosperms to other seed plants remains controversial<sup>1</sup>, great progress has been made in identifying the earliest extant splits in flowering-plant phylogeny, with the discovery that the New Caledonian shrub *Amborella trichopoda*, the water lilies (Nymphaeales), and the woody Austrobaileyales constitute a basal grade of lines that diverged before the main radiation in the clade<sup>2–8</sup>. By focusing attention on these ancient lines, this finding has re-written our understanding of angiosperm structural and reproductive biology, physiology, ecology and taxonomy<sup>9–12</sup>. The discovery of a new basal lineage would lead to further re-evaluation of the initial angiosperm radiation, but would also be unexpected, as nearly all of the ~460 flowering-plant families have been surveyed in molecular studies<sup>10</sup>. Here we show that Hydatellaceae, a small family of dwarf aquatics that were formerly interpreted as monocots, are instead a highly modified and previously unrecognized ancient lineage of angiosperms. Molecular phylogenetic analyses of multiple plastid genes and associated noncoding regions from the two genera of Hydatellaceae identify this overlooked family as the sister group of Nymphaeales. This surprising result is further corroborated by evidence from the nuclear gene phytochrome C (*PHYC*), and by numerous morphological characters. This indicates that water lilies are part of a larger lineage that evolved more extreme and diverse modifications for life in an aquatic habitat than previously recognized.

Molecular evidence has been particularly useful in clarifying the phylogenetic positions of groups with highly modified morphologies, such as holoparasites, mycoheterotrophs and certain aquatics. One such group that has resisted placement is Hydatellaceae (Fig. 1), an aquatic family of two genera (*Hydatella* and *Trithuria*), which is restricted to Australasia and India. These genera were traditionally included in Centrolepidaceae, a family of highly reduced monocots. The minute reproductive structures of both families are interpreted as multi-flowered inflorescences surrounded by bracts; individual flowers are unisexual, consisting of a single carpel or stamen, with no associated perianth or bract. Despite these similarities, numerous structural differences warranted separation of Hydatellaceae from Centrolepidaceae<sup>13</sup>. In fact, many features in Hydatellaceae are unknown among graminoid Poales (a clade consisting of Centrolepidaceae, grasses and five other families<sup>14</sup>), including monosulcate pollen, completely anatropous ovules and abundant starchy perisperm (seed storage tissue of nucellar origin). Other characteristics are virtually unknown in monocots; for example, cellular endosperm development, which is restricted to the basal monocot genus *Acorus*. A placement of Hydatellaceae within any of the major clades of

monocots has therefore been viewed as problematic<sup>15,16</sup>. Nonetheless, evidence from the plastid gene *rbcl* seemed to confirm conventional views by placing the one species studied (*Trithuria submersa*) in Poales<sup>17,18</sup>.

We show here that Hydatellaceae do not belong in Poales, or even in monocots, but instead diverged near the base of angiosperm phylogeny. We re-examined patterns of variation along the published *rbcl* sequence from *T. submersa* as part of a large-scale phylogenetic survey of grasses and relatives, and discovered that it is probably a PCR-based artefact representing a fusion product between a grass and a moss sequence (confirmed by J. Davis, Cornell University, personal communication). Subsequently, we found that combined analysis of multiple plastid genes from *T. submersa* identified it as the sister group of the water lilies (Nymphaeales), with strong bootstrap support from maximum parsimony and maximum likelihood analyses (Fig. 2).



**Figure 1 | Hydatellaceae.** **a**, Growth habit of *Trithuria submersa* showing a flowering individual with multiple inflorescence axes (each ~10–35 mm long). **b**, An inflorescence with several staminate flowers and ~10–20 pistillate flowers. The inflorescence is surrounded by an involucre of bracts. **c**, A pistillate flower with several stigmatic hairs, each consisting of a single row of cells<sup>16</sup>. Drawing by N. Oram, reprinted from ref. 31, with permission.

<sup>1</sup>UBC Botanical Garden and Centre for Plant Research (Faculty of Land and Food Systems), Centre for Biodiversity Research, and Department of Botany, 2357 Main Mall, The University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada. <sup>2</sup>Section of Evolution and Ecology, University of California, Davis, California 95616, USA. <sup>3</sup>Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland. <sup>4</sup>The Arnold Arboretum, 22 Divinity Avenue, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>5</sup>Royal Botanic Gardens, Mrs Macquaries Road, Sydney, New South Wales 2000, Australia. <sup>†</sup>Present address: Research Division, Canadian Museum of Nature, P.O. Box 3443, Station D, Ottawa, Ontario K1P 6P4, Canada.

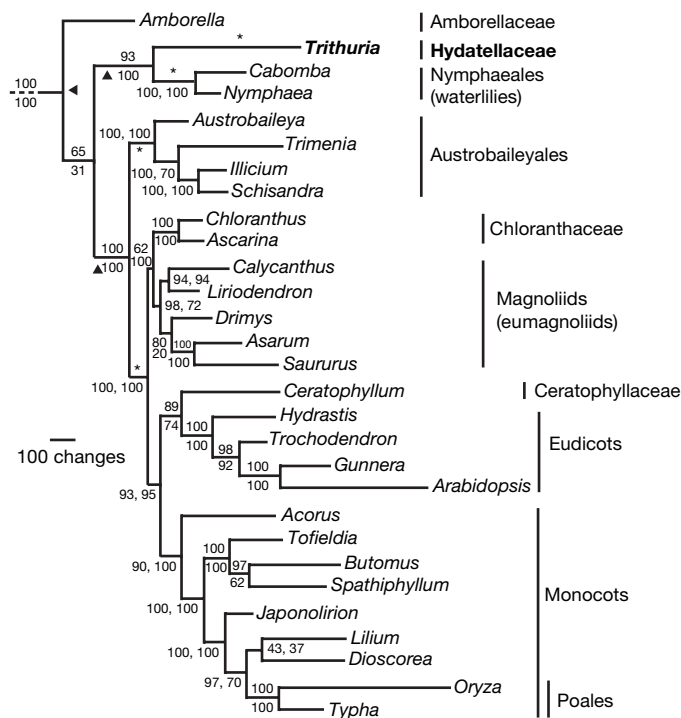
There is some uncertainty concerning the root of flowering-plant phylogeny<sup>2–8</sup>, probably a function of the relatively long branch connecting angiosperms to other seed plants. Because erroneous rooting can lead to mis-inference of ingroup relationships<sup>19</sup>, we determined plausible roots for the plastid multigene tree using the Shimodaira–Hasegawa test (see Supplementary Information). Hydatellaceae and Nymphaeales are sister groups in trees rooted at two positions not significantly worse than the optimal one (see arrowheads in Fig. 2), but a root on the *Trithuria* branch is among those rejected. The Hydatellaceae–Nymphaeales relationship is also recovered when substantially more exemplar taxa are included (see Supplementary Information). When each of six plastid data partitions (*atpB*; *ndhF*; *rbcL*; *rpl2*; the 3′-*rps12*–*trnL*(*caa*) region; and ten *psb* genes combined) is analysed in separate unrooted analyses (see Methods), we consistently observe a branch separating these two taxa from all other angiosperms, with strong maximum parsimony and maximum likelihood bootstrap support (90–100%).

To corroborate this result using evidence from another genome, we sampled a portion of the nuclear gene phytochrome C (*PHYC*) from *T. submersa*. When analysed with orthologous sequences from other angiosperms that include most of the lineages in Fig. 2, this gene also indicates that Hydatellaceae and water lilies are sister groups, with moderate to strong bootstrap support (70% and 90% in maximum parsimony and maximum likelihood analyses, respectively; Fig. 3a and Supplementary Information). We have also obtained plastid data from *Hydatella*, the other genus in the family, for two noncoding regions that span the plastid transfer RNA genes *trnL*(*uua*) and *trnF*(*gaa*) of *H. inconspicua*. We added these data to an

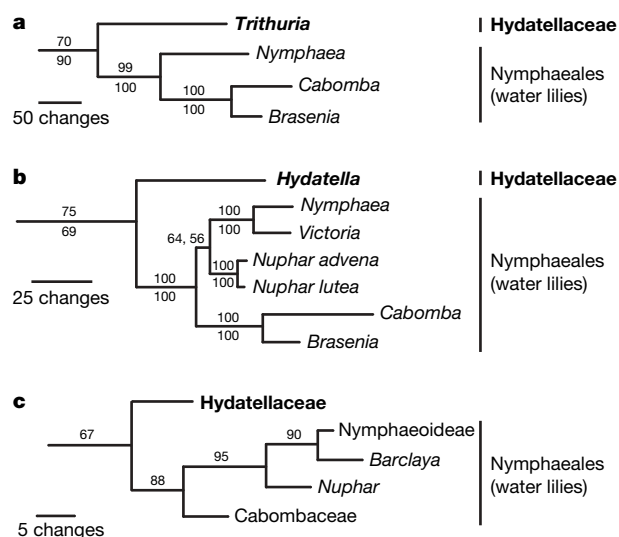
alignment for a somewhat larger plastid region surveyed across the major clades of angiosperms in Fig. 2 and several outgroup taxa<sup>6</sup>. Despite being based on a relatively limited amount of data, phylogenetic analyses again depict Hydatellaceae and water lilies as sister taxa, with moderate bootstrap support (75% and 69% from maximum parsimony and maximum likelihood bootstrap analyses, respectively; Fig. 3b).

Finally, we evaluated whether morphological evidence is consistent with these results by adding published data for Hydatellaceae and two graminoid Poales (Centrolepidaceae and Flagellariaceae) to a morphological matrix for basal angiosperms, including basal eudicots and monocots<sup>9,20</sup> (see Methods). Despite their extensive structural reduction, we could score Hydatellaceae for 64% of the characters in this matrix (see Supplementary Information). Analysis without Hydatellaceae gives some weakly supported results that are strongly over-ruled by molecular data (for example, Nymphaeales are linked with monocots and not located in the basal grade<sup>9</sup>). In most analyses we therefore constrained the arrangement of other taxa to a tree derived from recent molecular and combined morphological–molecular studies (see Methods), allowing the position of Hydatellaceae to be determined by morphology. However, it should be noted that when relationships are not constrained, Hydatellaceae and Nymphaeales form a clade with maximum parsimony bootstrap support comparable to that found with two of our molecular data sets (Fig. 3 and Supplementary Information).

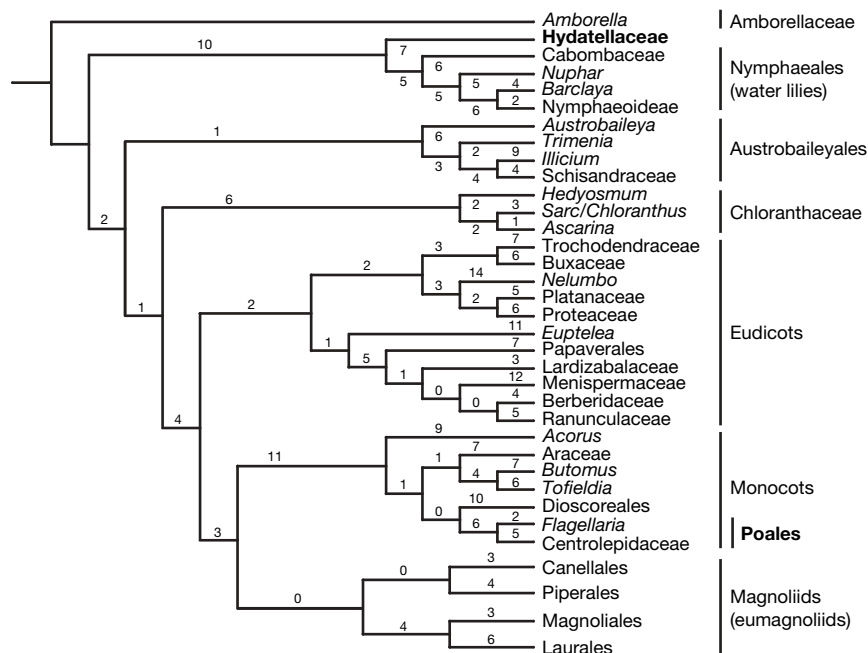
The constrained analysis also indicates that Hydatellaceae and Nymphaeales form a clade (Fig. 4), supported by ten unequivocal synapomorphies (lack of a vascular cambium, lack of pericyclic sclerenchyma, anomocytic stomata, truncate anther connective, boat-shaped pollen, inner integument with two cell layers, palisade exotesta, seed operculum formed by cell enlargement in the inner integument, perisperm and hypogeal germination). Some of these features are among those originally used to segregate Hydatellaceae from Centrolepidaceae<sup>13</sup>. Although most of them (except operculum) occur in other taxa, including various monocots, they are not consistently associated. Of eleven unequivocal synapomorphies of monocots as a



**Figure 2 | Phylogenetic placement of *Trithuria submersa* (Hydatellaceae) in angiosperms according to 17 plastid protein-coding loci and six associated noncoding regions.** Maximum parsimony and maximum likelihood analyses yield a topology (12,217 steps;  $-lnL = 86,798.648$ ) in which Hydatellaceae are sister to Nymphaeales. Branch lengths are maximum parsimony estimates (ACCTRAN optimization); bootstrap values are noted near branches (maximum parsimony above/left; maximum likelihood below/right); and outgroups (*Cycas* and *Ginkgo*) are trimmed for clarity (and the stem lineage shortened). The optimal and two suboptimal roots (arrowheads) cannot be distinguished from each other in the Shimodaira–Hasegawa test when considering these and four additional roots (asterisks, rejected at  $P = 0.001–0.048$ ). See Supplementary Information for species names.



**Figure 3 | Local placement of Hydatellaceae according to additional molecular data for *Hydatella* and *Trithuria*, and morphological data for the family as a whole.** All analyses support a sister-group relationship with Nymphaeales. **a**, The nuclear gene phytochrome C (*PHYC*) for *Trithuria submersa*. **b**, A portion of the plastid *trnT*–*F* region for *Hydatella inconspicua*. **c**, An unconstrained morphological analysis. Bootstrap support values are indicated (see Fig. 2); only parsimony bootstrap values are provided for morphology. The *PHYC* and morphological analyses include only angiosperms; the *trnT*–*F* analysis includes seed-plant outgroups. See Supplementary Information for full trees for angiosperms and species names.



**Figure 4 | Most parsimonious position of Hydatellaceae on the basis of morphology.** The arrangement of other taxa is constrained to correspond to angiosperm relationships found in recent molecular and combined morphological–molecular studies. Hydatellaceae are the sister group of

Nymphaeales. The numbers of unequivocal changes (with respect to the root indicated) are noted near each branch; tree length, 799 steps. For clarity, most phylogenetic structure within magnoliids has been excluded (see ref. 9 for details). *Sarc*, *Sarcandra*.

whole, only three occur in Hydatellaceae (no cambium, boat-shaped pollen and two-layered inner integument), and all of these also occur in Nymphaeales.

In the constrained analyses, a sister-group relationship of Hydatellaceae and Cabombaceae is two steps less parsimonious. The next-best position for Hydatellaceae, four steps less parsimonious than the optimal arrangement, is in monocots, as the sister group of Dioscoreales. Six other positions within monocots, one as the sister group of Centrolepidaceae, are an additional step less parsimonious. Two characters that would favour a relationship of Hydatellaceae with monocots rather than Nymphaeales are P2 sieve tube plastids (which arose independently within Aristolochiaceae) and linear leaves; although the exact point of origin of these characters is equivocal, they each arise twice on the shortest tree. The suboptimal association of Hydatellaceae with Centrolepidaceae would be supported by unisexual flowers, perianth loss and the single carpel, but this relationship would require reversals (or multiple origins) of other derived features of graminoid Poales, including globose, ulcerate pollen, orthotropous ovules and nuclear endosperm development.

Several other characters support a position of Hydatellaceae among the most basal angiosperms. These include completely ascidiate carpels and a four-nucleate embryo sac (the latter needs evaluation in the context of studies among related lineages<sup>11</sup>), features that are considered ancestral in angiosperms<sup>9,11</sup>; if Hydatellaceae and monocots were related, the presence of these characters in Hydatellaceae would represent reversals. Better information on other characters could affect morphological support for inferred relationships. For example, cotyledon number is unknown in Hydatellaceae, which have a 'minute, lens-shaped, incompletely developed embryo'<sup>16</sup>, as in some Nymphaeales.

Our results have little effect on previous inferences of the growth habit and ecology of the common ancestor of extant angiosperms<sup>12</sup>, because Hydatellaceae are so strongly linked with Nymphaeales, which are also aquatic herbs. A phylogenetic rooting of the flowering plants near *Amborella* and/or water lilies (Nymphaeales) has been robustly supported by numerous analyses of individual and combined sequences from the plastid and other genomes. One contrary result is probably a function of long-branch attraction and low taxon

sampling<sup>8</sup>. If *Amborella* alone, or a clade of *Amborella*, Nymphaeales and Hydatellaceae is sister to other angiosperms (see arrowheads in Fig. 2), parsimony optimization implies that the first angiosperms were woody and terrestrial, with Hydatellaceae and Nymphaeales an early line that invaded aquatic habitats. The third plausible rooting, with Hydatellaceae and Nymphaeales sister to all other angiosperms, would imply that the first angiosperms were either woody and terrestrial, or herbaceous and aquatic<sup>12</sup>.

It would be misleading to view Hydatellaceae merely as reduced water lilies. First, no member of Nymphaeales approaches the minute, submergence-tolerant, moss-like habit of Hydatellaceae. *Hydatella inconspicua* can grow and flower at >1 m depths; in other species plants are usually initially submerged, but may flower under water or on drying mud at the edges of seasonal pools or swamps. Second, Hydatellaceae have inflorescences rather than solitary flowers, indicating that their common ancestor with Nymphaeales could have had either condition. The possibility that Hydatellaceae are related to the Early Cretaceous aquatic *Archaeofructus*<sup>21</sup> should be investigated, as both taxa have inflorescences of naked, unisexual flowers.

Our current knowledge of Hydatellaceae is limited. The family has only recently been discovered in India, with speculation that it has been overlooked elsewhere<sup>16</sup>. Half of the 10 or so species were described in the past 25 yr, and there is substantial morphological variation among them; indeed, more species may await discovery. We lack information on phylogenetic relationships within the family, and it is clear that developmental morphology and monographic work should be high priorities, in addition to studies of the aquatic ecology, biogeography and conservation status of these curious and overlooked basal angiosperms.

## METHODS

**Molecular analyses.** Methods of DNA extraction, amplification, sequencing and alignment follow refs 3 and 22 for plastid data and ref. 4 for *PHYC*. Two sources of DNA were used to generate different *Trithuria submersa* regions: one prepared by the Royal Botanic Gardens, Kew (voucher: J. G. Conran 961 & P. J. Rudall, ADU), the other by K. Bremer, Uppsala University (voucher: Doust *et al.* 1123, MELU). We generated two new *rbcl* sequences using both sources; these are identical for their 483 base pair shared portion (we used the longer of the two (DQ915188, generated from the Conran and Rudall collection) in analyses here;



the shorter *rbcL* sequence (DQ915187) was generated from the Doust *et al.* collection). The problematic published *rbcL* for *T. submersa* (AF458076; refs 14, 17, 18 and 23) was obtained using a different DNA extraction generated from the Doust *et al.* collection; this sequence was not used here. We generated the *trnL-trnF* sequence for *Hydatella inconspicua* from cultivated material sourced from Lake Rotokawau, New Zealand (voucher: P. D. Chapman s.n., NSW accession 428712). All other source details are presented in the Nexus-formatted text files for each data set.

For the plastid-based analyses involving *T. submersa*, we combined protein-coding genes involved in a range of functions, including photosynthesis (*atpB*, *psb* (photosystem II) genes and *rbcL*), chlororespiration (*ndhF*, *ndhB*) and translation (*rpl2*, *rps7*, *rps12*). This multigene plastid matrix includes several conservative noncoding regions (that is, introns in *ndhB*, *rpl2*, and 3'-*rps12*; the intergenic spacers between these genes<sup>5</sup>, and between *ndhB* and *trnL*(*caa*)<sup>23</sup>), but we excluded more rapidly evolving spacer regions in two of the photosystem II gene clusters (*psbB-psbH* and *psbE-psbJ*) from consideration. For analyses of individual regions (subpartitions *atpB*; *ndhF*; *rbcL*; *rpl2*; 3'-*rps12-trnL*(*caa*) and the combined photosystem II genes), we excluded outgroups (*Cycas* and *Ginkgo*) to minimize the effect of rooting uncertainty on inferred bootstrap support values, as we were interested primarily in the structure of the angiosperm subtree in these analyses. We added the *trnL*(*uaa*)-*trnF*(*gaa*) region from *H. inconspicua* to a larger published matrix<sup>6</sup> for *trnT*(*ugu*)-*trnL*(*uaa*)-*trnF*(*gaa*), excluding several mutational hotspots from analysis<sup>6</sup>. We excluded one variable region in the *PHYC* alignment. We performed heuristic maximum-parsimony and maximum-likelihood searches using PAUP\* version 4.0b10 (ref. 24) and PHYML version 2.4.4 (ref. 25), respectively, using settings described in the Supplementary Information.

**Morphological analyses.** For morphological analyses, we added consensus of *Hydatella* and *Trithuria*, Centrolepidaceae, and Flagellariaceae to a published data matrix<sup>9,20</sup>, with modifications described in Supplementary Information; scoring of most characters was based on refs 13, 15, 16, 18 and 26–28, with additional data from primary sources provided in Supplementary Information. All heuristic maximum parsimony searches were performed using PAUP\*. We used two backbone constraint trees to fix relationships of all taxa but Hydatellaceae: one corresponding to the maximum parsimony tree from the combined morphological and molecular analyses of refs 9 and 20, the second (see Fig. 4) with modifications on the basis of increasingly robust molecular evidence on relationships in (eu)magnoliids<sup>5,7</sup> (Piperales, Canellales, Magnoliales, Laurales), eudicots<sup>29</sup> and monocots<sup>14,22,23</sup> (see also Fig. 2). The results reported above are based on the second backbone, but we found similar results using the first topology<sup>9,20</sup>. We also evaluated less than optimal arrangements of Hydatellaceae by searching for trees up to six steps less parsimonious than the optimal tree length, and by moving taxa manually with MacClade v. 4.06 (ref. 30). Heuristic maximum parsimony searches (including bootstrap analysis of the unconstrained data) otherwise followed settings described for the molecular analyses. We used the 'Trace all changes' tool in MacClade to determine unequivocal synapomorphies.

Received 13 September 2006; accepted 18 January 2007.

- Burleigh, J. G. & Mathews, S. Phylogenetic signal in nucleotide data from seed plants: implications for resolving the seed plant tree of life. *Am. J. Bot.* **91**, 1599–1613 (2004).
- Soltis, P. S., Soltis, D. E. & Chase, M. W. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402**, 402–404 (1999).
- Graham, S. W. & Olmstead, R. G. Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. *Am. J. Bot.* **87**, 1712–1730 (2000).
- Mathews, S. & Donoghue, M. J. Basal angiosperm phylogeny inferred from duplicate phytochromes A and C. *Int. J. Plant Sci.* **161** (6 Suppl.), S41–S55 (2000).
- Zanis, M. J., Soltis, D. E., Soltis, P. S., Mathews, S. & Donoghue, M. J. The root of angiosperms revisited. *Proc. Natl Acad. Sci. USA* **99**, 6848–6853 (2002).
- Borsch, T. *et al.* Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *J. Evol. Biol.* **16**, 558–576 (2003).
- Qiu, Y.-L. *et al.* Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Plant Sci.* **166**, 815–842 (2005).
- Leebens-Mack, J. *et al.* Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein zone. *Mol. Biol. Evol.* **22**, 1948–1963 (2005).
- Doyle, J. A. & Endress, P. K. Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. *Int. J. Plant Sci.* **161** (6 Suppl.), S121–S153 (2000).
- Angiosperm Phylogeny Group (APG II). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. *Bot. J. Linn. Soc.* **141**, 399–436 (2003).
- Williams, J. H. & Friedman, W. E. The four-celled female gametophyte of *Illicium* (Illiciaceae; Austrobaileyales): implications for understanding the origin and early evolution of monocots, eumagnoliids, and eudicots. *Am. J. Bot.* **91**, 332–351 (2004).

- Feild, T. S., Arens, N. C., Doyle, J. A., Dawson, T. E. & Donoghue, M. J. Dark and disturbed: a new image of early angiosperm ecology. *Paleobiology* **30**, 82–107 (2004).
- Hamann, U. Hydatellaceae—a new family of Monocotyledoneae. *N. Zeal. J. Bot.* **14**, 193–196 (1976).
- Bremer, K. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* **56**, 1374–1387 (2002).
- Dahlgren, R. M. T., Clifford, H. T. & Yeo, P. F. *The Families of the Monocotyledons: Structure, Evolution, and Taxonomy* (Springer, Berlin, 1985).
- Hamann, U. in *The Families and Genera of Vascular Plants IV. Flowering Plants. Monocotyledons. Alismatanae and Commelinanae (except Gramineae)* (ed. Kubitzki, K.) 231–234 (Springer, Berlin, 1998).
- Stevenson, D. W. *et al.* in *Monocots: Systematics and Evolution* (eds Wilson, K. L. & Morrison, D. A.) 17–24 (CSIRO, Collingwood, Australia, 2000).
- Michelangeli, F. A., Davis, J. I. & Stevenson, D. W. Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *Am. J. Bot.* **90**, 93–106 (2003).
- Graham, S. W., Olmstead, R. G. & Barrett, S. C. H. Rooting phylogenetic trees with distant outgroups: a case study from the commelinoid monocots. *Mol. Biol. Evol.* **19**, 1769–1781 (2002).
- Doyle, J. A. Early evolution of angiosperm pollen as inferred from molecular and morphological phylogenetic analyses. *Grana* **44**, 227–251 (2005).
- Sun, G. *et al.* Archaeofractaceae, a new basal angiosperm family. *Science* **296**, 899–904 (2002).
- Graham, S. W. *et al.* in *Monocots: Comparative Biology and Evolution (excluding Poales)* (eds Columbus, J. T., Friar, E. A., Porter, J. M., Prince, L. M. & Simpson, M. G.) 3–21 (Rancho Santa Ana Botanic Garden, Claremont, California, 2006).
- Chase, M. W. *et al.* in *Monocots: Comparative Biology and Evolution (excluding Poales)* (eds Columbus, J. T., Friar, E. A., Porter, J. M., Prince, L. M. & Simpson, M. G.) 63–75 (Rancho Santa Ana Botanic Garden, Claremont, California, 2006).
- Swofford, D. L. *Phylogenetic Analysis Using Parsimony\* (PAUP\*)* (Sinauer Associates, Sunderland, Massachusetts, 2002).
- Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
- Hamann, U. Neue Untersuchungen zur Embryologie und Systematik der Centrolepidaceae. *Bot. Jahrb. Syst.* **96**, 154–191 (1975).
- Cooke, D. A. in *The Families and Genera of Vascular Plants IV. Flowering Plants. Monocotyledons. Alismatanae and Commelinanae (except Gramineae)* (ed. Kubitzki, K.) 106–109 (Springer-Verlag, Berlin, 1998).
- Appel, O. & Bayer, C. in *The Families and Genera of Vascular Plants IV. Flowering Plants. Monocotyledons. Alismatanae and Commelinanae (except Gramineae)* (ed. Kubitzki, K.) 208–211 (Springer, Berlin, 1998).
- Kim, S., Soltis, D. E., Soltis, P. S., Zanis, M. J. & Suh, Y. Phylogenetic relationships among early-diverging eudicots based on four genes: were the eudicots ancestrally woody? *Mol. Phylog. Evol.* **31**, 16–30 (2004).
- Maddison, D. R. & Maddison, W. P. *MacClade 4: Analysis of Phylogeny and Character Evolution, Version 4.03* (Sinauer Associates, Sunderland, Massachusetts, 2001).
- Harden, G. J. (ed.) *Flora of New South Wales*. Vol. 4 (Univ. of New South Wales, Kensington, New South Wales, Australia, 1993).

Supplementary Information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are grateful to K. Bremer (Uppsala University) and the Royal Botanic Gardens, Kew, for generously providing DNAs, and to J. Conran, J. Davis, A. Doust, P. Rudall and D. Stevenson and other workers responsible for making the field collections and generating cultivated material. We acknowledge critical review of the manuscript by S. C. H. Barrett, M. W. Chase, T. S. Feild and E. M. Friis. This research was supported by an NSERC Discovery Grant to S.W.G., NSERC postgraduate scholarships to J.M.S. and H.S.R., Alberta Ingenuity and University Graduate Fellowship (University of British Columbia) funding to J.M.S., an NSF grant to S.M., and Royal Botanic Gardens Trust Sydney funding to A.D.M. and B.G.B.

**Author Contributions** Plastid data were generated by J.M.S., H.S.R. and A.D.M.; nuclear data were generated by S.M.; morphological data were compiled and scored by J.A.D., P.K.E. and B.G.B. Analyses were conceived and performed by S.W.G., J.A.D., J.M.S., H.S.R. and S.M. All authors contributed to the writing, which was coordinated by S.W.G. and J.A.D.

**Author Information** Novel sequences for this study have GenBank accession numbers as follows: *Aphelia brizula* (EF153935, EF153937, EF153939, EF153942, EF153945, EF153948, EF153950, EF153952, EF153954); *Brasenia schreberi* (DQ981792); *Centrolepis monogyna* (EF153934, EF153936, EF153938, EF153941, EF153944, EF153947, EF153949, EF153951, EF153953); *Hydatella inconspicua* (DQ916291); *Trithuria submersa* (AJ419142, DQ915185–DQ915189, DQ981794, EF153940, EF153943, EF153946); *Schisandra sphenanthera* (DQ981793). Alignments used are available for download in Supplementary Information. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W.G. ([swgraham@interchange.ubc.ca](mailto:swgraham@interchange.ubc.ca)).